Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

Network Theory
Lecture 4.01

EEU45C09 / EEP55C09
Self Organising Technological Networks

CONNECT
Networks of the Future

Nicola Marchetti
nicola.marchetti@tcd.ie

What is a network?

Profound Opening Quote:

A network is, in its simplest form, a collection of points joined together in pairs by lines. [1]

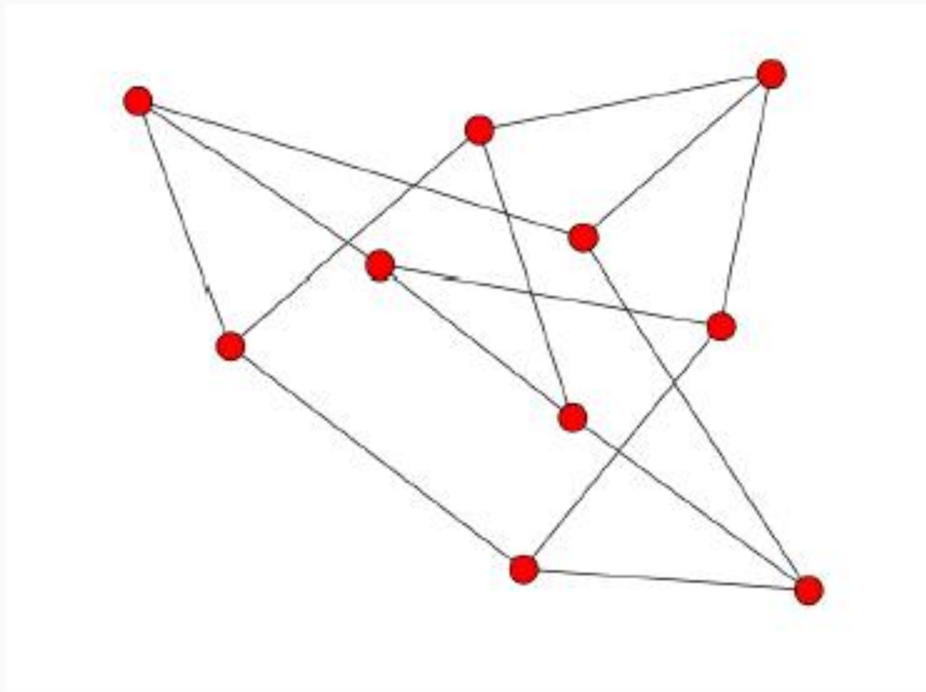


Figure 1: A graph representation of a network.

Types of Networks

Empirical Networks:

- Movie Actors
- Power Grid
- Food Webs
- World Wide Web
- Internet
- Yeast protein-protein interaction networks

Network Models:

- Random networks: Erdős-Rényi Network
- Small-World networks: Watts-Strogatz Model
- Scale-Free networks: Barabási-Albert Model

Empirical Networks

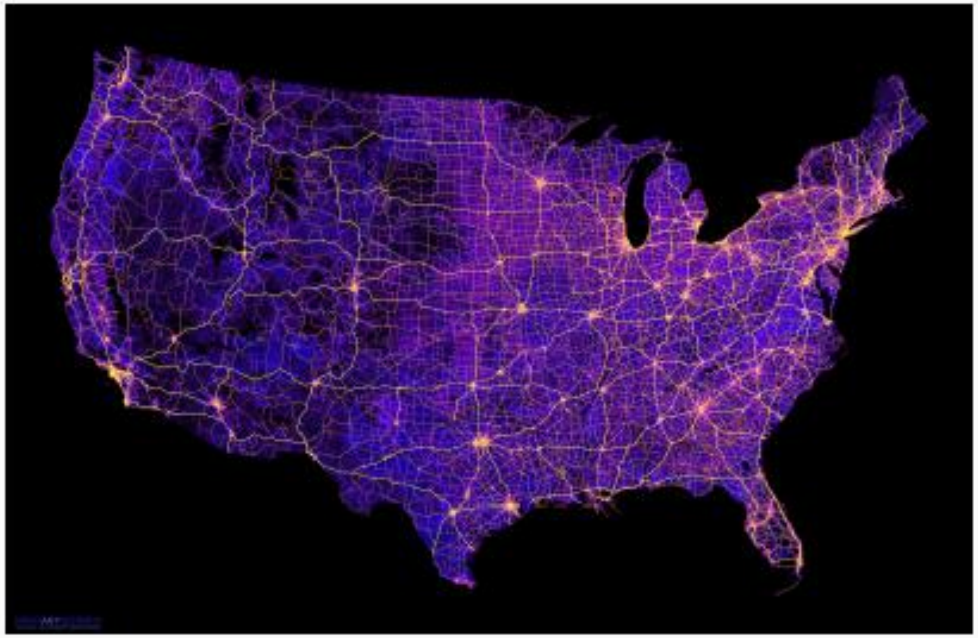


Figure 2: The US mapped exclusively using 8 million miles roads, streets and highways.

Source: robbibt
`//www.reddit.com/r/dataisbeautiful/duplicates/58gjl6/the_united_states_mapped_only_by_8_million_miles/`

Empirical Networks

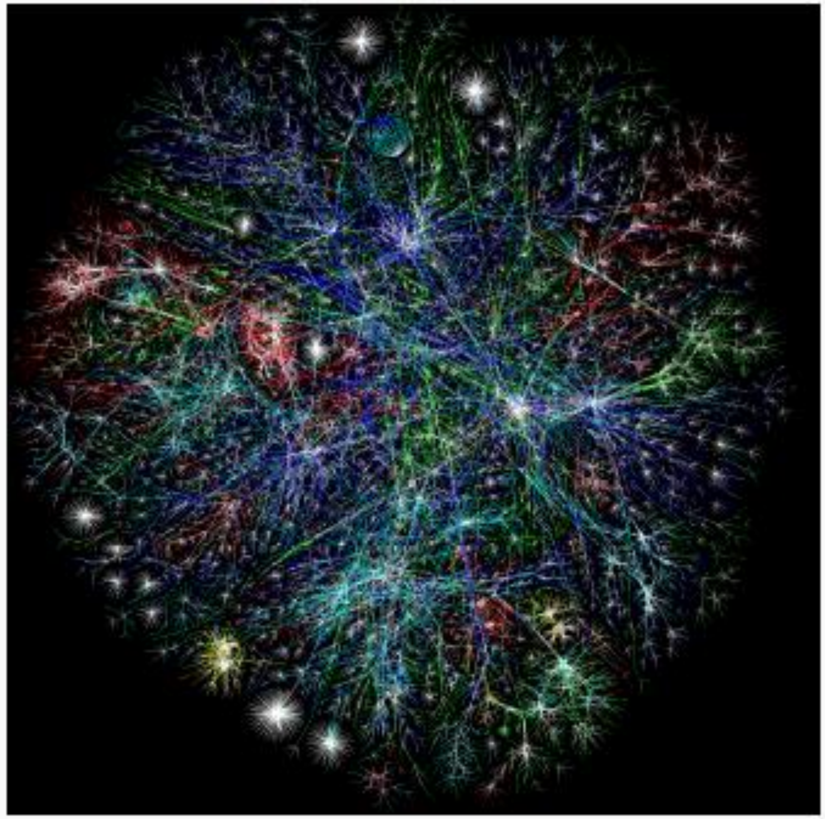
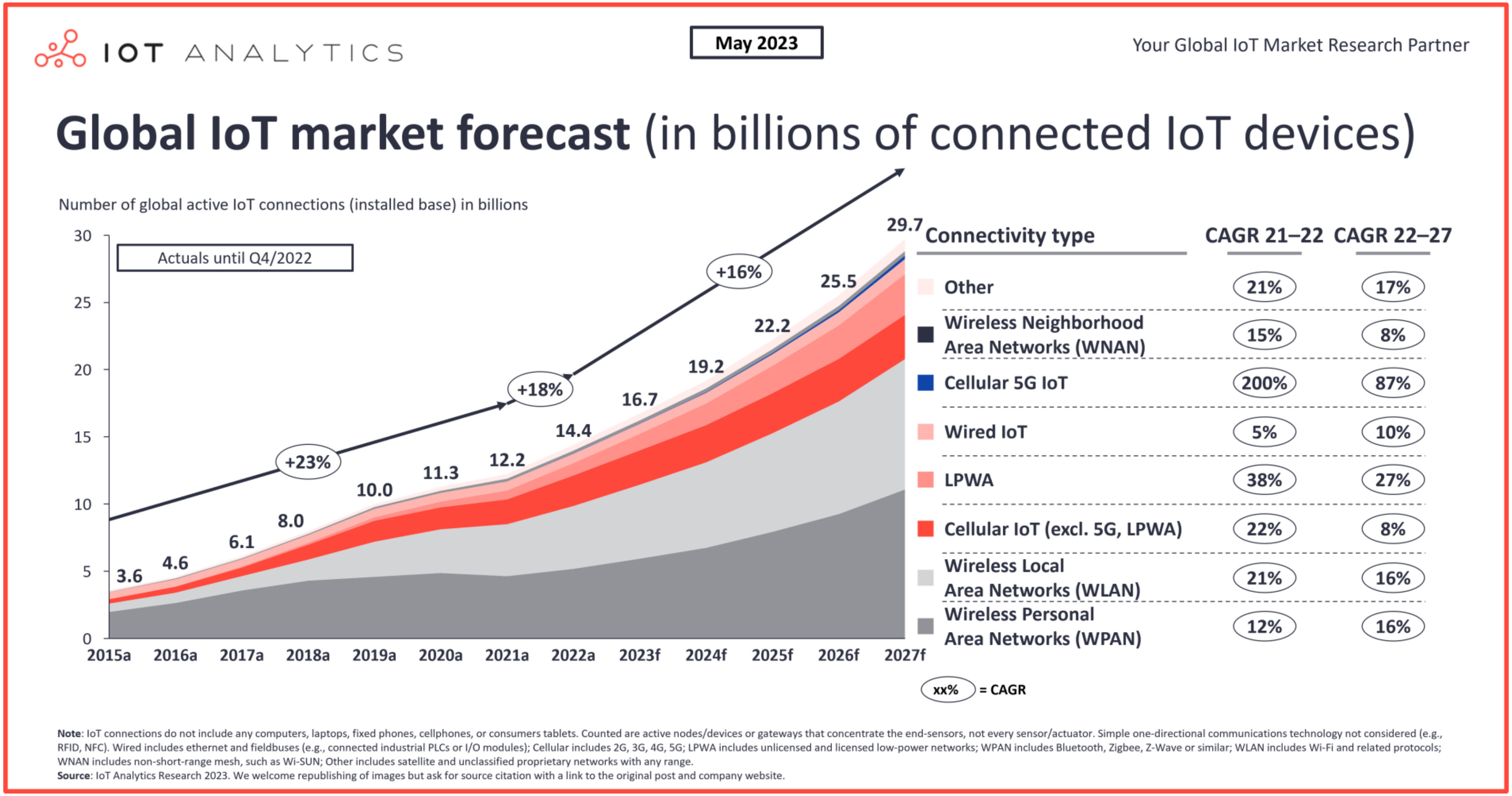


Figure 3: A map of the full internet (2003).

Source: The Opte Project

`http://blyon.com/blyon-cdn/opte/maps/`

How large are these networks?



https://iot-analytics.com/number-connected-iot-devices/

What questions can we ask?

- Do these networks display similar traits?
- Are networks created or are they emergent?
- Can simple rules lead to complex formations?

What insights does network science provide?

- Are these networks resistant to failure?
- Are bottlenecks common?
- How far are nodes from each other?
- Is there a path between two nodes?
- Can we “engineer” networks?

Network Science

Wikipedia:

"an academic field which studies complex networks such as telecommunication networks, computer networks, biological networks, cognitive and semantic networks, and social networks, considering distinct elements or actors represented by nodes (or vertices) and the connections between the elements or actors as links (or edges)." [3]

US National Research Council:

"the study of network representations of physical, biological, and social phenomena leading to predictive models of these phenomena." [4]

Lots of Interest and Tools

Interdisciplinary Academic Field:

- Graph Theory from Mathematics
- Statistical Mechanics from Physics
- Data Mining and Visualisation from Computer Science
- Inferential Modelling from Statistics
- Social Structure from Sociology

Properties of Interest

To address these topics, we need some more formality!

- Formation
- Structure
- Evolution
- Degeneracy
- Dynamics on networks: cooperation/competitivity, disease/information spreading

Graphs: A Formal Definition

Graph theory is the language of network science.

- A graph consists of a set of nodes (V) and edges (E).
- Edges can be directed (in one direction) or undirected (symmetric).

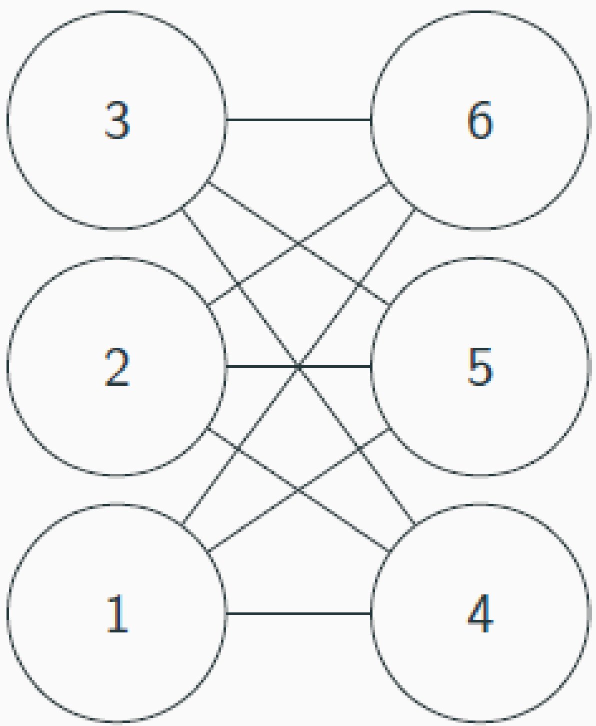
Figure 4: A simple graph (no self-edges or repeated edges) with two nodes and one undirected edge.

Graph Properties

- Network Size ($|V|$ and $|E|$)
- Degree (k): number of neighbours
- Average Degree ($\langle k \rangle$)
- Degree Distribution ($P(k)$): probability of a given degree
- Path Length (l_{ab}): shortest number of hops between nodes a and b
- Average Path Length ($\langle l \rangle$)
- Clustering Coefficient (C)
- Centrality
- More advanced (not today!): Spectral methods, Path problems, Graph ensembles

Adjacency Matrix (A_{ij})

$$A_{ij} = \begin{cases} 1, & \text{if there is an edge between vertices } i \text{ and } j. \\ 0, & \text{otherwise.} \end{cases}$$



(a) Utility Graph

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

(b) Adjacency matrix (A_{ij}) of Utility Graph

Centrality

Which are the most important or central vertices in a network?

The degree, otherwise known as the degree centrality is one simple measure of this.

Degree (k)

The degree of each node i can be found using the adjacency matrix:

$$k_i = \sum_{j=1}^{|V|} A_{ij}$$

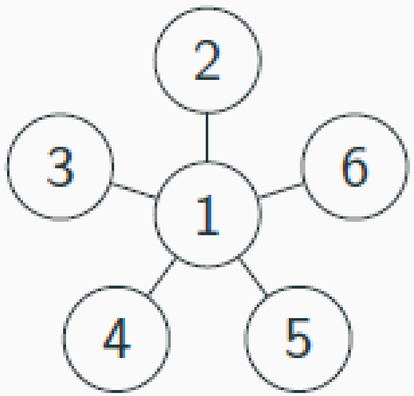


Figure 5: The Star graph with 6 nodes. What is the degree of each node?

Degree Distribution ($P(k)$)

The degree distribution $P(k)$ is the number of nodes with a given degree k.

$$P(k) = \sum_{i=1}^{|V|} \delta_{k_i,k} / \, |V|$$

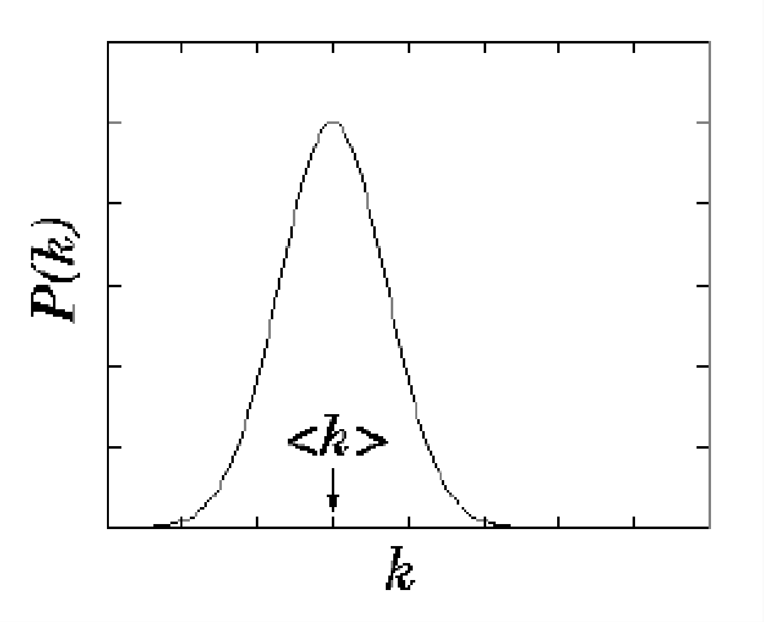


Figure 6: Degree distribution and average degree of a Poisson distributed random graph.

Path Length (l)

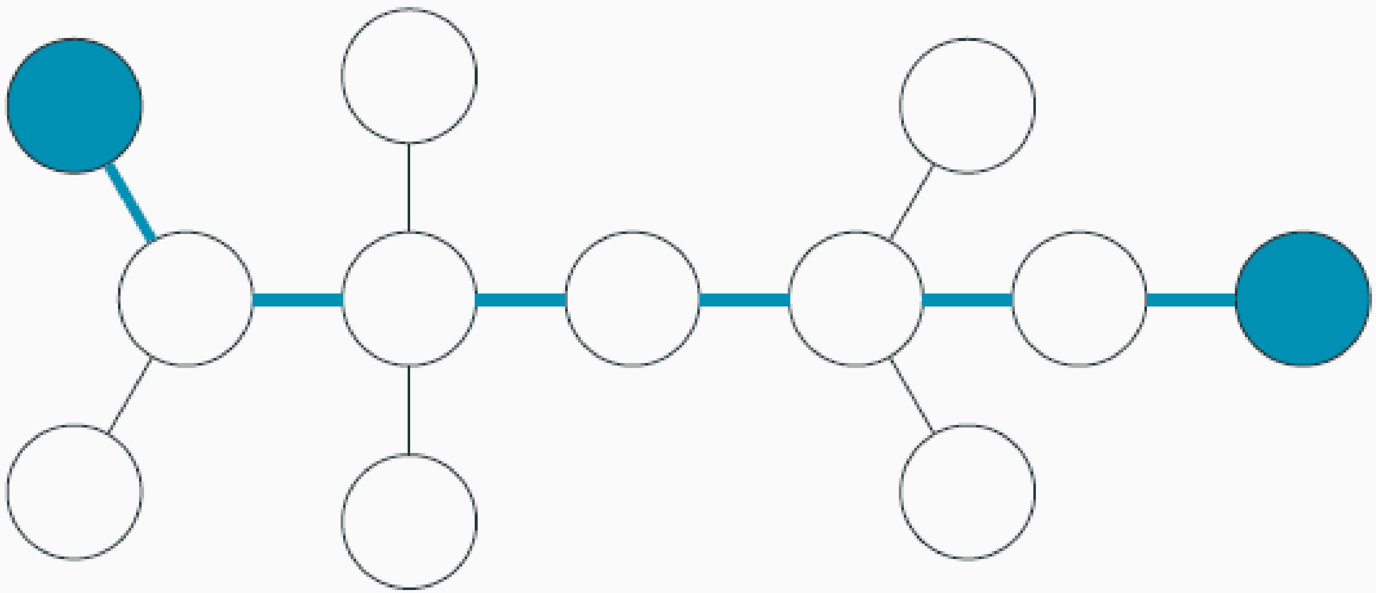
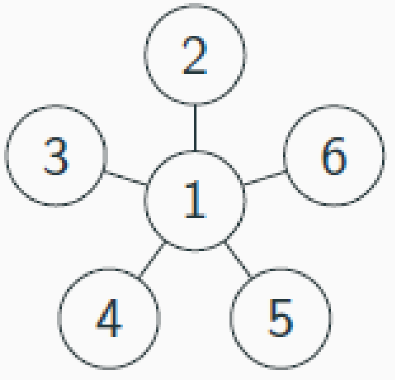


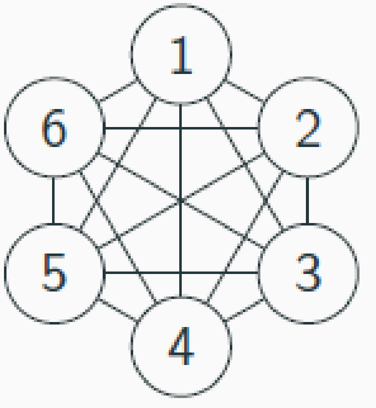
Figure 7: Trees have a unique path between each node. The path length between nodes trivially becomes the number of edges separating these two nodes.

Clustering Coefficient (C)

Local clustering coefficient (c_i) is the fraction of neighbours of node i which are interlinked.



(a) Local $c_1 = 0$ for node 1



(b) Local $c_1 = 1$ for node 1

The overall clustering coefficient,

$$C = \frac{1}{|V|} \sum_{i=1}^{|V|} c_i .$$

Random Graphs

Networks with a complex topology and unknown organising principles often appear random; thus random-graph theory is regularly used in the study of complex networks. [5]

Random Graphs

ER Random Graphs

Erdős and Rényi define a random graph as N labelled nodes connected by n edges, which are chosen from the $N(N-1)/2$ possible edges.

Probability space of random graphs

There is a set of $\binom{N(N-1)/2}{n}$ graphs, each with N nodes and n edges, with each graph having equal probability. (Where $\binom{a}{b}$ is the binomial coefficient.)

Random Graphs

We will see that random graphs have a short average path length but a low clustering coefficient (therefore bad for modelling social networks etc).

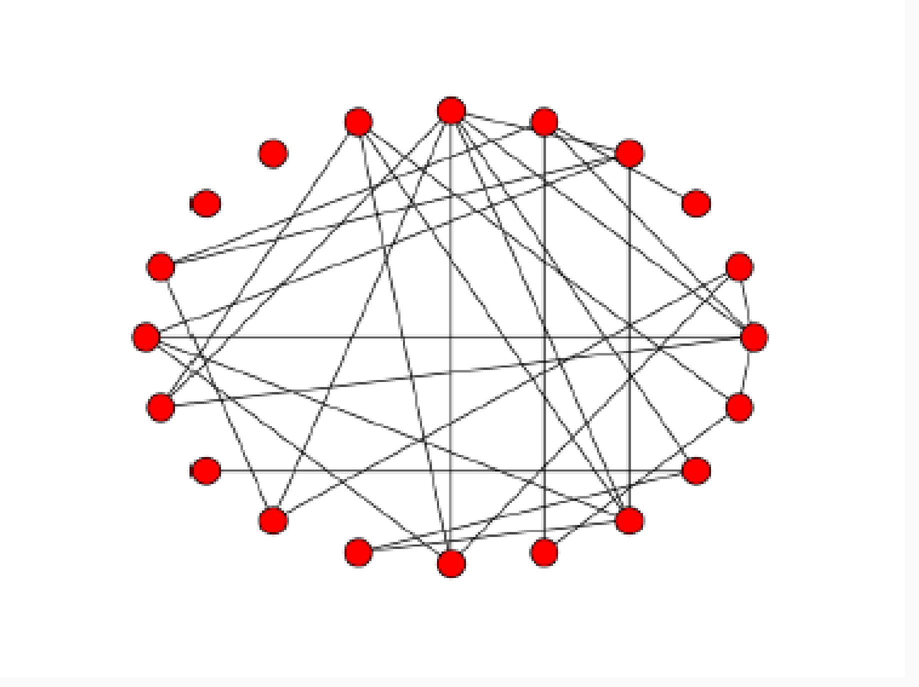


Figure 8: Erdős-Rényi model [6]: $|V| = 20$, $p = 0.15$

Building Random Graphs

The construction of a random graph is called an evolution:

For $i=1,\dots,N-1$

- Starting with N isolated vertices, consider v_i
- With probability p, independently construct an edge between v_i and every other vertex with a higher vertex label v_j, where $j > i$.

Random Graphs: Average number of edges ($\langle n \rangle$)

Due to p, the total number of edges is a random variable. The average number of edges is,

$$\langle n \rangle = p \left(N(N-1)/2 \right).$$

The probability of creating a graph G_0 with N nodes and n edges is,

$$P(G_0) = p^n (1-p)^{N(N-1)/2-n},$$

where p is the probability of an edge being created and $(1-p)$ is the probability of it not being created.

Random Graphs: Degree distribution

Random graphs occur when each node randomly chooses to connect to the $N-1$ other nodes independently. We find a collection of graphs with the degree distribution,

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k},$$

which in the large graph limit can be approximated by,

$$P(k) \approx \frac{(Np)^k \exp[-Np]}{k!} = \exp[-\langle k \rangle] \frac{\langle k \rangle^k}{k!}.$$

Random Graphs: Average degree ($\langle k \rangle$)

The average degree in a random graph is found using the definition,

$$
\begin{aligned}
\langle k \rangle &= \sum_{k=0}^{N-1} k \cdot P(k), \\
&= \sum_{k=0}^{N-1} k \cdot \binom{N-1}{k} p^k (1-p)^{N-1-k}, \\
&= \ldots, \\
&= (N-1)p
\end{aligned}
$$

Random Graphs: Diameter (d)

The diameter of a graph is the maximal distance between any pair of its nodes.

The number of nodes at a distance l is not much smaller than $\langle k \rangle^l$. When all nodes are within this distance, we can say that

$$\langle k \rangle^d \sim N,$$

where d is the diameter of the graph.

Random Graphs: Diameter (d)

Therefore, we can find an expression for the diameter,

$$d \log \langle k \rangle \sim \log N,$$
$$d \sim \frac{\log N}{\log \langle k \rangle}.$$

We can perform a similar estimate for the average path length to find,

$$l \sim \frac{\log N}{\log \langle k \rangle}.$$

Random Graphs: Clustering Coefficient (C)

The clustering coefficient concerns how many neighbours of a given node are themselves connected. In a random graph, the probability of a node and its nearest neighbour being connected is the same as any two nodes being connected.

Therefore,

$$C_{random} = p = \frac{\langle k \rangle}{N}.$$

For random graphs, the clustering coefficient falls linearly with number of nodes N. Real-world networks do not become less clustered as the network grows.

- M. Newman, *Networks: An Introduction*.
  New York, NY, USA: Oxford University Press, Inc., 2010.
- A. Sangiovanni-Vincentelli, "Let's get physical: adding physical dimensions to cyber systems," *Internet of Everything Summit, Rome*, 2014.
- Wikipedia, "Network science — wikipedia, the free encyclopedia," 2016.
  [Online; accessed 12-December-2016].
- T. C. on Network Science for Future Army Applications, *Network Science*.
  National Research Council, 2006.
- R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, pp. 47–97, Jan 2002.
- P. Erdős and A. Rényi, "On random graphs," *Publicationes Mathematicae Debrecen*, vol. 6, pp. 290–297, 1959.

Ackowledgement

- Dr Neal McBride, Arista Networks